AN APPLICATION OF THREE TESTS FOR NONADDITIVITY

BU-343-M* W. T. Federer October, 1970

Abstract

Three tests for nonadditivity were applied to a numerical example involving two categories of classification to illustrate the procedures. The example contained six treatments in six blocks of a randomized complete blocks design and the observation was number of leatherjackets present on a given plot or area. One test used is John Tukey's test for nonadditivity wherein the expected response is proportional to the product of the row and column effects. A second test used is appropriate when the expected value of the response is the square of the sum of the overall mean, the row effect, and the column effect. The third test used is one making use of the contingency table chi-square deviations. The sum of squares associated with each of the tests was approximately the same with Tukey's test being associated with the largest sum of squares for this particular example.

*Mimeo Series, Biometrics Unit, Cornell University.

AN APPLICATION OF THREE TESTS FOR NONADDITIVITY

BU-343-M* W. T. Federer October, 1970

An application of three tests for nonadditivity to a numerical example involving a two-way classification was made in order to illustrate the procedures and to obtain a comparison of the tests. One of the tests for nonadditivity is the well-known test given by Tukey [1949] which is especially suitable if the expected response is proportional to the product of row and column effects in a two-way classification. The second test for nonadditivity is one developed by D. S. Robson which is appropriate when the expected value of the response is the square of the sum of the row and column effects. The third test presented is one involving the chi-square contingency table model (Federer [1959]).

The example used for the illustrations and comparisons is the one given in Bartlett's [1947] paper. The example is presented in Table 1 and involves v=6 treatments (two are controls (untreated) and four represent treatments to control leatherjackets) in r=6 blocks of a randomized complete blocks design. The yields Y_{ij} represent counts of leatherjackets in a given area or plot. (Additional examples involving controls and treated plots may be found in Bartlett [1936,1947] and Beall [1942].)

The analysis of variance for the data in Table 1 is presented in Table 2. Three treatment contrasts (control 1 vs. control 2, controls vs. treated, and among the four different treatments for leatherjacket control) are given in Table 2; the individual blocks $\times$ treatment contrasts are partitioned and utilized in computing the F values presented in the table. The apparent variance inequality does not allow use of the pooled remainder variance equal to 288.02.

*Mimeo Series, Biometrics Unit, Cornell University.

Table 1. Leatherjacket counts

	Treatments						Totals	Means	Dev.
	Controls		Treated						
	1	2	3	4	5	6	$Y_{.j}$	$\bar{y}_{.j}$	$(\bar{y}_{.j}-\bar{y})$
Block I	92	66	19	29	16	25	247	41.17	148/36
II	60	46	35	10	11	5	167	27.83	-332/36
III	46	81	17	22	16	9	191	31.83	-188/36
IV	120	59	43	13	10	2	247	41.17	148/36
V	49	64	25	24	8	7	177	29.50	-272/36
VI	134	60	52	20	28	11	305	60.83	496/36
Totals $Y_{i.}$	501	376	191	118	89	59	1334	-	0
means $\bar{y}_{i.}$	83.50	62.67	31.83	19.67	14.83	9.83	-	37.06	-
dev. $36(\bar{y}_{i.}-\bar{y})$	1672	922	-188	-626	-800	-980	0	-	-

Table 2. Analysis of variance of leatherjacket counts

Source of variation	d.f.	Sum of squares	Mean square	F	F_{05}
Total	36	85256	-		
Correction for mean	1	49432.11	-		
Blocks	5	2358.22	-		
Treatments	5	26265.22	-		
Control 1 vs 2	1	1302.08	-	1.46	6.61
Controls vs treated	1	23364.01	-	81.81	6.61
Among treated	3	1599.03	533.01	6.08	3.29
Remainder	25	7200.45	288.02		
Controls × blocks	5	4457.42	891.48		
Controls vs tr. × bl.	5	1427.90	285.58		
Treated × blocks	15	1315.22	87.68		
Tukey's NA	1	2147.27	-		
Residual	24	5053.18	210.54		
Robson's NA	1	1857.58	-		
Residual	24	5342.87	222.62		
Contingency NA	1	2027.58	-		
Residual	24	5172.87	215.54		

Prior to computing the sum of squares for Tukey's one-degree-of-freedom for nonadditivity, the nature of a multiplicative response is considered. In a two-way row X column classification if the expected response of the observation Y_{ij} is $\mu\rho_i\gamma_j$ where ρ_i are row effects, γ_j are column effects, and μ is a constant, the row means then are $\sum_{j=1}^{c}\mu\rho_i\gamma_j/c = \mu\rho_i\bar{\gamma}_.$; the column means are $\sum_{i=1}^{r}\mu\rho_i\gamma_j/r = \mu\gamma_j\sum_{i=1}^{r}\rho_i/r = \mu\gamma_j\bar{\rho}_.$. The overall mean is $\mu\bar{\gamma}_.\bar{\rho}_.$. In an errorless model then, the deviation or interaction effect is $\mu\rho_i\gamma_j - \mu\rho_i\bar{\gamma}_. - \mu\gamma_j\bar{\rho}_. + \mu\bar{\rho}_.\bar{\gamma}_. = \mu(\rho_i-\bar{\rho}_.)(\gamma_i-\bar{\gamma}_.)$. The estimate of this quantity will be proportional to $(\bar{y}_{i.}-\bar{y})(\bar{y}_{.j}-\bar{y}) = e'_{ij}$. The sum of squares for the regression of the deviation for the additive model, $\hat{e}_{ij} = Y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}$, on the e'_{ij} is Tukey's one-degree-of-freedom for nonadditivity. This sum of squares is computed as:

$$\left(\sum_{i=1}^{r}\sum_{j=1}^{c}\hat{e}_{ij}e'_{ij}\right)^2 \Big/ \sum_{i=1}^{r}\sum_{j=1}^{c}(e'_{ij})^2$$

$$= \frac{[\sum_i\sum_j (Y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y})(\bar{y}_{i.}-\bar{y})(\bar{y}_{.j}-\bar{y})]^2}{\sum_i(\bar{y}_{i.}-\bar{y})^2\sum_j(\bar{y}_{.j}-\bar{y})^2}$$

$$= \frac{[\sum_i\sum_j Y_{ij}(\bar{y}_{i.}-\bar{y})(\bar{y}_{.j}-\bar{y})]^2}{\sum_i(\bar{y}_{i.}-\bar{y})^2\sum_j(\bar{y}_{.j}-\bar{y})^2}$$

which is equal to

$$(60782.02)^2/1{,}720{,}532.4 = 2147.27\ .$$

It should be noted here that the logarithmic transformation makes the row and column effects additive in this case.

For the test proposed by D. S. Robson, the row and column effects are additive under the square root transformation. The expected value of the observation Y_{ij} is $(\mu+\rho_i+\gamma_j)^2 = \mu^2 + \rho_i^2 + \gamma_j^2 + 2\mu\rho_i + 2\mu\gamma_j + 2\rho_i\gamma_j$. The column mean is computed as $\sum_{i=1}^{r}(\mu+\rho_i+\gamma_j)^2/r = \mu^2 + \overline{\rho_{.}^2} + \gamma_j^2 + 2\mu\bar{\rho}_{.} + 2\mu\gamma_j + 2\bar{\rho}_{.}\gamma_j$; the row mean is $\mu^2 + \rho_i^2 + \overline{\gamma_{.}^2} + 2\mu\rho_i + 2\mu\bar{\gamma}_{.} + 2\rho_i\bar{\gamma}_{.}$; the overall mean is $\mu^2 + \overline{\rho_{.}^2} + \overline{\gamma_{.}^2} + 2\mu\bar{\rho}_{.} + 2\mu\bar{\gamma}_{.} + 2\bar{\rho}_{.}\bar{\gamma}_{.}$. Therefore, Y_{ij} - row mean - column mean + overall mean is equal to $2(\rho_i-\bar{\rho}_{.})(\gamma_j-\bar{\gamma}_{.})$. The estimated values for $(\gamma_j-\bar{\gamma}_{.})$ and $(\rho_i-\bar{\rho})$ are obtained from the equations:

$$\widehat{(\gamma_j-\bar{\gamma}_{.})} = \sqrt{\bar{y}_{.j}-\sigma_\rho^2} - \sum_{j=1}^{c}\sqrt{\bar{y}_{.j}-\sigma_\rho^2}\Big/ c$$

and

$$\widehat{(\rho_i-\bar{\rho}_{.})} = \sqrt{\bar{y}_{i.}-\sigma_\gamma^2} - \sum_{i=1}^{r}\sqrt{\bar{y}_{i.}-\sigma_\gamma^2}\Big/ r \quad .$$

where

$$\sigma_\rho^2 = \sum_{i=1}^{r}\widehat{(\rho_i-\bar{\rho}_{.})}^2/r \quad ,$$

$$\sigma_\gamma^2 = \sum_{j=1}^{c}\widehat{(\gamma_j-\bar{\gamma}_{.})}^2/c \quad ,$$

and σ_ρ^2 and σ_γ^2 satisfy the following:

$$\sum_{j=1}^{c}\sqrt{\bar{y}_{.j}-\sigma_\rho^2}\Big/ c = \sum_{i=1}^{r}\sqrt{\bar{y}_{i.}-\sigma_\gamma^2}\Big/ r \quad .$$

Computationally, the simple method of solving for $\widehat{(\rho_i-\bar{\rho}_{.})}$ and $\widehat{(\gamma_j-\bar{\gamma}_{.})}$ is to first set either σ_ρ^2 or σ_γ^2 equal to zero and proceed iteratively. The convergence is rather rapid. Let $e^{+}_{ij} = 2(\widehat{\rho_i - \bar{\rho}_.})(\widehat{\gamma_j - \bar{\gamma}_.})$. These residuals are computed in Table 3. Then $(\sum_i \sum_j \hat{e}_{ij} e^{+}_{ij})^2 / \sum_i \sum_j (e^{+}_{ij})^2$ is a one-degree-of-freedom sum of squares for nonadditivity. For the data of Table 1 and from the deviations listed in Table 3, this sum of squares is equal to $(796.2713)^2/341.3303 = 1857.58$ as listed in Table 1.

In the contingency chi-square table the errorless model would have the observation $Y_{ij} = \mu_{i.}\mu_{.j}/\mu$, where $\mu_{i.}$ is the row mean, $\mu_{.j}$ is the column mean, and μ is the overall mean. Replacing the preceding quantities by the row mean $\bar{y}_{i.}$, the column mean $\bar{y}_{.j}$, and the overall mean $\bar{y}$, the sum of squares due to observed minus calculated is:

$$\sum_{i=1}^{r} \sum_{j=1}^{c} \left(Y_{ij} - \frac{Y_{i.}Y_{.j}}{Y_{..}}\right)^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} (Y_{ij} - \bar{y}_{i.}\bar{y}_{.j}/\bar{y})^2 \quad .$$

The sum of squares for the additive model is $\sum_i \sum_j (Y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y})^2 = \sum_i \sum_j \hat{e}^2_{ij}$. Therefore, the difference in the two sums of squares, i.e., RCA $= \sum_i \sum_j \hat{e}^2{}_{ij} - \sum_i \sum_j (Y_{ij} - Y_{i.}Y_{.j}/Y_{..})^2$, is a measure of nonadditivity. For the above example this difference is $7200.45 - 5172.87 = 2057.58$. This sum of squares is almost as large as the one obtained from Tukey's one-degree-of-freedom sum of squares. As indicated by Federer [1959] this sum of squares can be negative. In this case the absolute value of the difference was utilized. Since it is not known if the remaining sum of squares is orthogonal to the above difference, since the sum of squares can be negative, and since the model can be described as a form of the multiplicative model (as illustrated below), the above procedure has not been pursued. Since the deviation in the errorless model is $\mu_{i.}\mu_{.j}/\mu - \mu_{i.} - \mu_{.j} + \mu$ and is estimated by $\bar{y}_{i.}\bar{y}_{.j}/\bar{y} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y} = \frac{1}{\bar{y}}(\bar{y}_{i.} - \bar{y})(\bar{y}_{.j} - \bar{y}) = e^{-}_{ij}$, a sum of squares similar to Tukey's and Robson's above would be

Table 3. Residuals for four models

		Treatment					
		1	2	3	4	5	6
Block I	$\hat{e}_{1j}$	4.39	- 0.78	- 16.94	5.22	- 2.94	11.06
	e'_{1j}	190.94	105.29	- 21.47	- 71.49	- 91.36	-111.91
	e^{+}_{1j}	2.74	1.78	- 0.03	- 0.99	- 1.46	- 2.04
	e^{-}_{1j}	5.15	2.84	- 0.58	- 1.93	- 2.46	- 3.02
Block II	$\hat{e}_{2j}$	- 14.28	- 7.44	12.39	- 0.44	5.39	4.39
	e'_{2j}	-428.32	-236.19	48.16	160.36	204.94	251.05
	e^{+}_{2j}	- 5.81	- 3.76	0.06	2.10	3.09	4.31
	e^{-}_{2j}	- 11.56	- 6.37	1.30	4.33	5.53	6.78
Block III	$\hat{e}_{3j}$	- 32.28	23.56	- 9.61	7.56	6.39	4.39
	e'_{3j}	-242.54	-133.75	27.27	90.81	116.05	142.16
	e^{+}_{3j}	- 3.03	- 1.96	0.03	1.10	1.61	2.25
	e^{-}_{3j}	- 6.55	- 3.61	0.74	2.45	3.13	3.84
Block IV	$\hat{e}_{4j}$	32.39	- 7.78	7.06	- 10.78	- 8.94	- 11.94
	e'_{4j}	190.94	105.29	- 21.47	- 71.49	- 91.36	-111.91
	e^{+}_{4j}	2.74	1.78	- 0.03	- 0.99	- 1.46	- 2.04
	e^{-}_{4j}	5.15	2.84	- 0.58	- 1.93	- 2.46	- 3.02
Block V	$\hat{e}_{5j}$	- 26.94	8.89	0.72	11.89	0.72	4.72
	e'_{5j}	-350.91	-193.51	39.46	131.38	167.90	205.68
	e^{+}_{5j}	- 4.63	- 3.00	0.05	1.67	2.46	3.44
	e^{-}_{5j}	- 9.47	- 5.22	1.06	3.55	4.53	5.55
Block VI	$\hat{e}_{6j}$	36.72	- 16.44	6.39	- 13.44	- 0.61	- 12.61
	e'_{6j}	639.90	352.86	- 71.95	-239.58	-306.17	-375.06
	e^{+}_{6j}	7.98	5.17	- 0.09	- 2.89	- 4.25	- 5.93
	e^{-}_{6j}	17.27	9.52	- 1.94	- 6.46	- 8.26	- 10.12

$(\sum_i \sum_j \hat{e}_{ij} e^-_{ij})^2/\sum_i \sum_j (e^-_{ij})^2$ which for the above example is $(1640.2133)^2/1252.3205 = 2147.3129$; this is identical to Tukey's sum of squares. This is what one would expect since $e^-_{ij} = e'_{ij}/\bar{y}$. Thus for this method of computing the sums of squares the two methods are identical, i.e., $(\sum_i \sum_j \hat{e}_{ij} e'_{ij})^2/\sum_i \sum_j (e'_{ij})^2 = (\sum_i \sum_j (\hat{e}_{ij} e^-_{ij})^2/\sum_i \sum_j (e^-_{ij})^2$. Also, the models are identical in that for the multiplicative model the row mean is $\sum_{j=1}^{c} \mu\rho_i\gamma_j/c = \mu\rho_i\bar{\gamma}_.$, the column mean is $\mu\bar{\rho}_.\gamma_j$, and the overall mean is $\mu\bar{\rho}_.\bar{\gamma}_.$. Therefore, the (row mean) $\times$ (column mean) divided by the overall mean as in the chi-square contingency table is $(\mu\rho_i\bar{\gamma}_.)(\mu\bar{\rho}_.\gamma_j)/\mu\bar{\rho}_.\bar{\gamma}_. = \mu\rho_i\gamma_j$ which is the multiplicative model.

It is not known if the quantity with the largest sum of squares indicates the transformation most nearly making the data additive. In the above example it would appear that the logarithmic transformation would be slightly superior to the square root transformation with regard to additivity. These data are counts and hence might be expected to have a multinomial distribution for which the contingency table chi-square would be appropriate; this would again indicate the logarithmic transformation to make the multiplicative model an additive one. However, one might believe that these counts followed a Poisson distribution and might make a square root transformation <u>to stabilize the variances</u>; the square root transformation might not produce effects which are additive. Likewise, the transformation making the effects additive need not necessarily stabilize the variances.

Bartlett [1947] suggested that these data be transformed to square roots of counts prior to calculating the analysis of variance. When one calculates the analysis of variance on the counts and on the square root of counts for the treatment contrasts given in Table 3, some interesting results are obtained using the pooled estimate of the error variance in both cases. These results are:

Source of variation	d.f.	F values		
		Y	$\sqrt{Y}$	tabulated (.05)
Blocks	5	1.6	2.0	2.60
Control 1 vs Control 2	1	4.52	2.6	4.24
Controls vs others	1	81.12	109.9	4.24
Among treated	1	1.85	5.4	2.99

If one were making significance statements at the five per cent level, quite different statements would be made for Y and for $\sqrt{Y}$. However, if one utilized the individual variances for the contrasts as was done in Table 2, no difference in statements would have been made.

Literature Cited

Bartlett, M. S. [1936]. The square root transformation in the analysis of variance. J. Royal Statistical Soc., Supplement 3:68-78.

Bartlett, M. S. [1947]. The use of transformations. Biometrics 3:39-52.

Beall, G. [1942]. The transformation from entomological field experiments so that the analysis of variance becomes applicable. Biometrika 32:243-262.

Federer, W. T. [1959]. A note on additivity. Biometrics Unit Mimeo Series No. BU-101-M, Cornell University.

Robson, D. S. [1970]. Tests for non-additivity viewed as tests of the hypothesis of no interaction. A preliminary report. BU-334-M in the Biometrics Unit Mimeo Series, Cornell University.

Tukey, J. [1949]. One degree of freedom for non-additivity. Biometrics 5:232-242.